



An Epidemic Model of Violence and Public Support in Civil War

Citation

Zhukov, Yuri Maximovich. 2013. An epidemic model of violence and public support in civil war. *Conflict Management and Peace Science* 30(1): 24–52.

Published Version

doi:10.1177/0738894212456936

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:11878768>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

An Epidemic Model of Violence and Public Support in Civil War

Yuri M. Zhukov*

Forthcoming in *Conflict Management and Peace Science*

January 12, 2012

Abstract

How do civilians respond to violence in civil war, and how do these responses shape combatants' coercive strategies? Conventional wisdom expects civilian victimization to backfire, as a security-minded public “balances” against the side posing the greatest threat to its livelihood and survival. Yet combatants often expect a terrorized population to do the opposite, “bandwagoning” with those most willing and capable to inflict harm. Using an epidemic model of popular support dynamics, I explore the logic of balancing and bandwagoning in irregular civil war. I argue that when civilian strategy is clearly communicated to combatants, civilians are always better off balancing, and combatants are better off avoiding punishment. When civilian choice is not observed, the balancing equilibrium breaks down and patterns of violence depend on the local balance of power. The model's results challenge the view that selective violence is most common in areas of incomplete control. Due to uncertainty over civilian behavior, violence in both divided and perfectly-controlled areas can occur in equilibrium, inflicting great costs on civilians. I compare these predictions against the historical record of Soviet counterinsurgency in Western Ukraine, using new micro-level data from the declassified archives of the Soviet secret police.

*Department of Government, Harvard University. I am grateful to Robert Bates, William Bossert, Jeff Friedman, Andrew Radin, Brandon Stewart, Dustin Tingley, three anonymous reviewers and seminar participants at MIT, Yale, Harvard and the Association for the Study of Nationalities for helpful conversations and feedback on earlier versions of this model. All remaining errors are my own.

How do civilians respond to violence in civil war, and how do these responses shape combatants' coercive strategies? Irregular intrastate war typically involves a violent competition for the support of the population.¹ To the side able to secure it, popular support facilitates the extraction of provisions, tributes and taxes, and generates a supply of military recruits, administrative personnel and informants. The larger this pool of resources, the easier it becomes to sustain military operations and build the institutions of a sovereign state. For civilians, however, choosing sides is risky business. Cooperation with insurgents invites punishment by the government, and cooperation with the government invites punishment by insurgents.² Combatants make strategic choices based on expectations of how civilians will respond to this punishment: by "balancing" against the side that inflicts the most costs, or by "bandwagoning" with it. When civilians balance, the violent interaction becomes a race to the bottom: victory will go to whichever side can minimize civilian costs. When civilians bandwagon, the interaction becomes a race to the top: victory will go to the side willing to hurt civilians the most. In this sense, bandwagoning is inefficient: it creates incentives for the escalation of punishment, increasing the human and material costs of war. If security-seeking civilians are always better off by balancing, why would combatants ever risk driving a terrorized population into the arms of the enemy?

Using an epidemic model of popular support dynamics, I explore the logic of balancing and bandwagoning in civil war, and the incentives these civilian strategies present for combatants. I argue that punishment can be avoided only when civilian strategy is readily observable. When civilian choice is uncertain, several patterns of violence are likely to emerge. Two-sided violence occurs where the initial balance of power is evenly divided. One-sided violence occurs where the balance is asymmetric.³ Where one side has dominant but incomplete control, violence by the weaker combatant is more likely. Where one side

¹An *irregular war* is defined as an organized armed conflict, characterized by a "dearth of large-scale direct military confrontations . . . and the absence of frontlines" (Kalyvas, 2005, 91). The centrality of the population to combatant strategy in irregular war is stated succinctly by the U.S. Army's Counterinsurgency Field Manual 3-24: "Popular support allows counterinsurgents to develop the intelligence necessary to identify and defeat insurgents" (Headquarters, Department of the Army, 2006, 1-29).

²*Punishment* is here defined as any use of selective (Kalyvas, 2006, 12) or direct (Bacells, 2011, 399-400) violence: individual or collective arrests or executions of suspected opponents, carried out with light weapons on the basis of private information provided by civilians. Although selective violence can be highly inaccurate in practice, depending on combatants' ability to identify opponents (Kalyvas, 2006, 189-191), this definition excludes generally indiscriminate forms of violence like area bombardment and cordon-and-search. *Cooperation* is defined as the provision of information or other forms of support (i.e. material, financial) to help locate a suspected opponent, or otherwise assist in her capture or execution (Kalyvas 2006, 174-176, Bacells 2011, 399-400).

³*Territorial control* (or *sovereignty*) is here defined as the ability to extract resources for war-making (e.g. information, tributes, taxes) from a population (Tilly, 1985, 181).

has fully consolidated control, violence by the stronger combatant is more likely. These propositions challenge the view that selective violence is most common in areas of incomplete control (Kalyvas, 2006, 2008b). Due to the uncertainty of civilian behavior, selective violence in both divided and perfectly-controlled areas can occur in equilibrium, high risks of collateral damage notwithstanding. I compare these predictions against the stylized facts of Soviet counterinsurgency in Western Ukraine, using new micro-level data from the declassified archives of the Soviet secret police.

Civilian victimization in war has been the subject of a growing volume of theoretical and empirical research, although deep divisions remain over whether such violence suppresses enemy support or inflames it. The conventional wisdom is that killing civilians is usually counterproductive (Pape 1996; Arreguín-Toft 2001; Francisco 2004; Abrahms 2006; Saxton and Benson 2008; Kocher et al. 2011, and also Christia 2008). Civilian targeting can compel an insecure public to withdraw its support and side with the opposition, “balancing” against the biggest threat to civilian survival. A classic example of this phenomenon can be found in Nazi-occupied Yugoslavia, where German reprisals against civilians alienated the local population and facilitated partisan recruitment (Hehn, 1979; Arreguín-Toft, 2003). More recently, retired General Stanley McChrystal articulated the balancing perspective in his May 2010 statement that U.S. forces in Afghanistan should exercise “courageous restraint” to avoid civilian casualties.⁴ If balancing is the dominant response to collateral damage, combatants keen on securing civilian cooperation should, as a rule, avoid escalating their use of punishment. If high rates of violence do take place, they will generally be a result of some miscalculation or erroneous assumption about civilian choice (Kalyvas, 2006, 162-65).

This view has been disputed by a parallel body of research on cases where repression and mass killing contributed to military success (Hibbs, 1973; Tilly, 1978; Stoll, 1993; Downes, 2006, 2007; Lyall, 2009, 2010). In this perspective, civilians can be compelled to support the biggest killer, “bandwagoning” with the side that shows itself willing and capable of inflicting great physical harm. The hanging courts and concentration camps used by the British during the Mau Mau uprising in Kenya are sometimes cited in this context: as crude, but effective deterrents against support for rebels (Peters, 2007). The bandwagoning perspective was also apparent in Muammar Qaddafi’s vow in February 2011 to track down and kill Libyan protesters “house by house,” reflecting the expectation that civilians can be terrorized into supporting the side most willing to hurt them.⁵ If bandwagoning is the dominant civilian

⁴Sean D. Naylor, “McChrystal: Civilian deaths endanger mission,” *Army Times*, 30 May 2010.

⁵Kareem Fahim and David D. Kirkpatrick, “Qaddafi’s Grip on the Capital Tightens as Revolt Grows,” *The New York Times*, 22 February 2011.

survival strategy, escalation is not necessarily inefficient. Indeed, combatants will resort to punishment precisely because it works.

The balancing-bandwagoning debate is far from settled: punishment often occurs in practice, and civilians respond with both types of strategies. The empirical variety of relationships between killing and popular support has prompted the development of theoretical models capable of accounting for both the deterrent and escalatory impacts of coercion (Lichbach, 1987; Moore, 1998, 2000; Carey, 2006). Most such efforts have been grounded in microeconomic producer theory, where actors seek an optimal allocation of resources between violence and non-violence (Lichbach, 1987; Moore, 1998). Others have given more formal consideration to the strategic interaction between incumbents and opponents (Crescenzi, 1999; Pierskala, 2010). Much recent theoretical work has focused on one-sided violence conducted by an incumbent regime in peacetime (Lskavyan, 2007; Gregory et al., 2011) or by the winning party to a civil war (Esteban et al., 2010), although less attention has been devoted to the competitive dynamics of two-sided violence.

The literature has yielded a wealth of useful insights into the determinants of compromise and violence, but important questions remain. First, under what conditions are civilians better off selecting a balancing strategy, and under what conditions is bandwagoning preferable? Few, if any, theoretical efforts have explicitly sought to accommodate both patterns of cooperation in a unified model. Second, conditional on the strategy of civilians, what is the optimal level of force each combatant should use? If, as the bandwagoning school contends, civilians can maximize their chances of survival by supporting the greater of two evils, the motivation behind punishment is straightforward: victory will go to whichever side can terrorize civilians the most. If, however, civilians are always better off by balancing, why should we ever observe the use of selective violence where the risk of collateral damage is high? After all, victory should go to whichever side can terrorize civilians the least.

Using an epidemic model of popular support dynamics, this paper offers two potential explanations for punishment: information problems and local asymmetries. In the first instance, combatants are unsure of how civilians will react to mounting casualties – by balancing against the side responsible for the majority of deaths, or by bandwagoning with it. To cope with this uncertainty, combatants alternate between low and high levels of force. In the second instance, punishment is used to exploit or compensate for local disadvantages in personnel, intelligence and recruitment.⁶ In places where initial conditions slightly favor

⁶This insight is broadly consistent with empirical findings from Valentino et al. (2004); Hultman (2007); Wood (2009) and Kocher et al. (2011).

one side, bandwagoning encourages the disadvantaged combatant to escalate. Where the advantage is overwhelming, one-sided violence by the hegemon occurs regardless of civilian strategy. Where conditions approach parity, incentives exist for violence on both sides.

The epidemic model lends itself fittingly to the study of civil conflict. While destructiveness and contagion are obvious substantive motivations to use infectious disease as an analogy for conflict, there are perhaps even more compelling theoretical and methodological reasons to move beyond the metaphor and formally adapt mathematical models of epidemics to the agenda of conflict research.⁷ First, the epidemic model places population dynamics at the center of the analysis, enabling the derivation of predictions about the flow of public support from one fighting side to another. Second, the model is inherently dynamic, offering insights not only into which equilibrium is reached, but also the process by which each equilibrium is reached. Third, the model offers a flexible, “workhorse” foundation for the study of war, capable of accommodating increasing layers of causal complexity.

Although traditional mathematical epidemiology takes behavioral choice to be exogenous, this paper considers optimizing strategic behavior on the part of the players.⁸ I employ an epidemic model of popular support dynamics as a payoff function for a simple three-player game, and derive the players’ best response strategies under several informational assumptions, in areas of contested and complete control. I also develop an evolutionary agent-based model to examine the adaptation of civilian and combatant strategies in repeated play. By closing the gap between epidemic modeling and game theory, the following paper accounts for some of the literature’s more puzzling findings and yields a number of novel predictions.

The paper is structured as follows. I begin with a simple narrative of civil conflict as a struggle for popular support, and formalize its logic and mechanisms with a system of ordinary differential equations. I then derive the model’s equilibria and discuss how their stability is influenced by rates of punishment and levels of territorial control. A simulation follows, in which three sets of actors – civilians, insurgents and the government – interact and select optimal strategies through an evolutionary process. The analytical and simulated results are then compared against stylized historical facts from the post-WWII Soviet counterinsurgency campaign in Western Ukraine, using new disaggregated data from the archives of the NKVD. The paper concludes with several summary remarks.

⁷Previous uses of the epidemic *analogy* in security studies and conflict research have included Siverson and Starr (1991); Buhaug and Gleditsch (2008); Braithwaite (2010). Previous efforts to formally apply epidemic *modeling* to the social sciences have included Epstein (1997); Epstein and Axtell (1996); Kolesin (1998, 2000, 2010); Zhukov (2012).

⁸Earlier works that have considered adaptive behavior in epidemiological modeling include Kremer (1996); Epstein et al. (2008).

1 The Narrative

Imagine a hypothetical conflict zone with three sets of actors: civilians, insurgents and the government. Sovereignty is contested between the government and insurgents, and the two combatants compete for the support of the civilian population. To the side able to secure it, popular support will bring taxes, manpower, food, supplies and intelligence. The extraction of these resources is essential to the military effort and, ultimately, to the establishment of a viable sovereign state (Elton, 1975; Tilly, 1985). Crucially, civilian cooperation is necessary for the identification of one's rivals and the production of direct, or selective, violence against them (Kalyvas, 2006; Balcells, 2010, 2011).

Sitting on the fence, civilians are interested in security above all else (Kalyvas, 2008b, 406). In deciding whether to support insurgents or the government, civilians seek to maximize their own chances of survival – by choosing the side that can most credibly provide protection (balancing), or by joining the side causing them the most harm (bandwagoning).⁹

The government and the insurgents want to entice civilians to cooperate with them and punish those who side with their opponents, but have differential opportunities to do so. In irregular war, the combatants' agents and collaborators tend to hide among the civilian population, creating an "identification problem" whose severity depends on the balance of territorial control (Kalyvas 2006, 89-91; Kalyvas 2008b, 407). Where the public already supports the insurgents, the active opposition becomes difficult for the government to identify. Wary of punishment by insurgents, locals are hesitant to provide intelligence. Because security forces are unable to correctly distinguish the insurgents' base of support from the peaceful population, civilians are arrested and disappeared along with combatants. Likewise, where the public supports the state, insurgents have difficulty exercising selective violence against government supporters. State presence, monitoring and retaliatory capacity are too pervasive for insurgents to identify defectors without error. In each case, violence remains selective by intent, but its targets are selected inaccurately in practice (Kalyvas, 2006, 189).

In this atmosphere of uncertainty, the two combatants must decide on an optimal level of force, while accounting for their inability to fully control how accurate their use of selective violence will be. Civilians – as the potential, if often unintentional, victims of selective violence – must choose balancing or bandwagoning as an optimal survival strategy. If they

⁹The logic of cooperation applies primarily to non-ethnic civil wars, and ethnic conflicts where members of one group are given the opportunity to defect to the other side. Where this is not the case (e.g. wars of extermination), this logic has limited traction and cooperation patterns resemble those in conventional war (Kalyvas, 2008a).

balance, they will cooperate at a higher rate with the side inflicting the least harm against civilians. If they bandwagon, they will cooperate at a higher rate with the side inflicting the most harm. This cycle of punishment and cooperation continues until an equilibrium is reached, in which either one side fully asserts control, or some form of stalemate is achieved.

An emerging conventional wisdom, most famously articulated by Kalyvas (2006, 111-112), is that both civilian cooperation and patterns of violence are endogenous to the balance of territorial control.¹⁰ Because civilians are most likely to cooperate where it is safe for them to do so, selective violence is expected primarily in areas where its perpetrator enjoys dominant, but incomplete territorial control – where the combatant’s presence is sufficient to facilitate civilian cooperation, but not so hegemonic as to completely cut off opponents’ access to the population. By contrast, where territorial control approaches parity and both sides are present in equal force – and thus equally capable of punishing those who help their enemies – selective violence is expected to be rare (Kalyvas, 2006, 204). Here, combatants lack the intelligence needed to avoid civilian deaths and abstain from violence to avoid encouraging support for their opponents. Where one side exercises complete territorial control, selective violence by the hegemon is unlikely because it is unnecessary, while the weaker side is isolated from the population and can only use indiscriminate force (Kalyvas, 2006, 220).

In sum, violence in partially-controlled areas is expected to be common; in divided and fully-controlled areas, it is off the equilibrium path. These predictions rest on the assumption that rational civilians balance rather than bandwagon: “everything else being equal, most people prefer to collaborate with the political actor that best guarantees their survival” (Kalyvas, 2008b, 406). Where selective violence is so inaccurate as to appear indiscriminate, it is avoided because it is counterproductive. As I show below, dropping this assumption allows us to explain why several other patterns of violence often occur: why insurgents in a position of weakness may terrorize civilians, why strong governments may freely repress opponents, why violence between two equally-matched combatants may spiral out of control.

The following section formalizes this stylized narrative and addresses two central questions. First, under what conditions are civilians better off selecting a balancing strategy, and under what conditions is bandwagoning preferable? Second, conditional on the strategy of civilians, what is the optimal level of punishment each combatant should apply?

¹⁰For additional research on the endogeneity of selective and indiscriminate violence in this context, see Mason and Krane (1989); Kalyvas (1999, 2004, 2008b); Goodwin (2001); Kalyvas and Kocher (2007); Kocher et al. (2011); Balcells (2011).

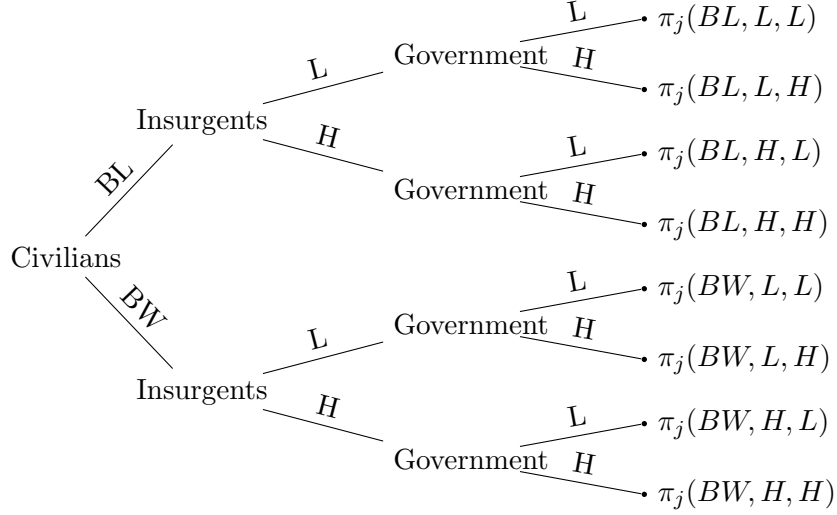


Figure 1: **Game Tree.** BL and BW represent civilian balancing and bandwagoning strategies. L and H represent low and high levels of insurgent or government violence. $\pi_j(\cdot)$ represents the payoffs associated with each strategy set for players $j \in \{C, I, G\}$.

2 The Dynamics of Popular Support

The conflict zone is populated by three sets of actors: civilians (C), insurgent supporters (I) and government supporters (G). These groups are assumed mutually exclusive and exhaustive, such that $C + I + G = N$ and $\frac{C}{N} + \frac{I}{N} + \frac{G}{N} = 1$, where N is the total size of the population. The civilians' objective is to stay alive. The insurgents' and government's objectives are to monopolize their respective levels of public support.

Each agent must select some strategy s to pursue her goals. Based on their assessment of the likelihood of survival in each case, civilians choose between strategies of balancing and bandwagoning, $s_C \in \{BL, BW\}$. Based on their assessment of expected public support in each case, insurgents and the government choose between low and high levels of violence, $s_I \in \{L, H\}$, $s_G \in \{L, H\}$. This choice set is shown graphically in Figure 1.

Payoffs for each actor are denoted by $\pi_C(\cdot)$, $\pi_I(\cdot)$ and $\pi_G(\cdot)$, and depend on the strategy chosen by each actor and a set of initial conditions, discussed at length below. Civilian payoffs are represented by the disutility of being punished by the two combatants, which includes the human and material costs directly or indirectly inflicted by the insurgents' and government's use of force. Ranging from -1 (highest costs) to 0 (no costs), these costs are highest when both combatants play H , lowest if both play L , and intermediate if one plays H and the other L . Insurgent and government payoffs are directly linked to the equilibrium

balance of active support for the two combatants. These payoffs range from 0 (no support) to 1 (complete support), with 1/2 representing an evenly split balance of support.

The decision problem for each actor is to choose a strategy that maximizes her payoffs, conditional on the choices available to the other two actors. Section 1 offered a brief narrative of how these strategies might generate varying levels of civilian casualties and public support. Section 2.1 introduces an epidemic model of popular support dynamics to formally specify the payoff function mapping the strategy space to these outcomes.

Table 1: **Notation Table**

Symbol	Description	Operationalization
Population parameters		
C_t	total neutral civilians at time t	$C_t \in [0, \infty)$
I_t	total insurgent supporters at time t	$I_t \in [0, \infty)$
G_t	total government supporters at time t	$G_t \in [0, \infty)$
Strategy choices		
β	civilian cooperation strategy (s_C)	$\beta \in \{0, 1\}$, with 1=balance, 0=bandwagon
ρ_I	insurgents' punishment strategy (s_I)	$\rho_I \in \{0, 1\}$, with 1=high, 0=low
ρ_G	government's punishment strategy (s_G)	$\rho_G \in \{0, 1\}$, with 1=high, 0=low
Exogenous parameters		
θ_I	insurgents' territorial control	$\theta_I = 1 - \theta_G \in [0, 1]$, with 1=full, 0=none
θ_G	government's territorial control	$\theta_G = 1 - \theta_I \in [0, 1]$, with 1=full, 0=none
k	constant civilian immigration rate	$k \in (0, \infty)$
u	constant population death rate	$u \in (0, \infty)$
Endogenous parameters		
μ_I	rate of civilian cooperation with insurgents	$\mu_I = \beta(1 - \theta_G)\rho_G + (1 - \beta)(1 - \theta_I)\rho_I$
μ_G	rate of civilian cooperation with government	$\mu_G = \beta(1 - \theta_I)\rho_I + (1 - \beta)(1 - \theta_G)\rho_G$
Payoff functions		
$\pi_C(\cdot)$	payoffs to civilians (costs of punishment)	$\pi_C(\cdot) = -((1 - \theta_I)\rho_I + (1 - \theta_G)\rho_G)$
$\pi_I(\cdot)$	payoffs to insurgents (equilibrium support)	$\pi_I(\cdot) = I_{eq}/(I_{eq} + G_{eq})$
$\pi_G(\cdot)$	payoffs to government (equilibrium support)	$\pi_G(\cdot) = G_{eq}/(I_{eq} + G_{eq})$

2.1 The Model

The rate of change in public support is modeled as a function of punishment, territorial control and cooperation. These dynamics are shown graphically in Figure 2, where C represents neutral civilians, I represents insurgent supporters, and G represents government supporters. Table 1 summarizes the various parameters, their symbology and operationalization.

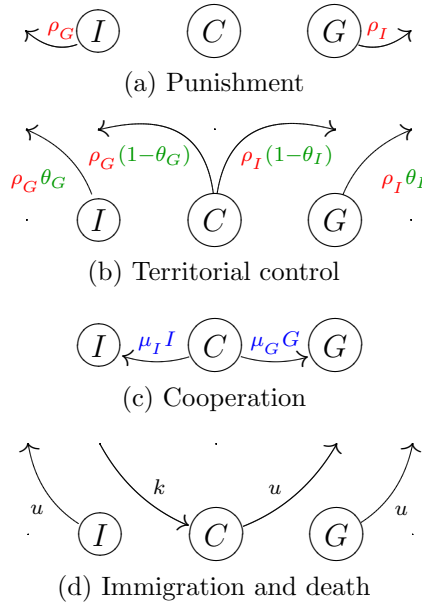


Figure 2: **Support flows.**

- a) **Punishment.** Insurgents and the government regulate the size of each other's groups through punishment – a label that includes any form of selective violence that forcibly removes combatants from the battlefield (Kalyvas, 2006, 173-74). The government punishes insurgent supporters at rate ρ_G and the insurgents punish government supporters at rate ρ_I (Figure 2a). For the government, this rate can be interpreted as the number of arrests, executions, deportations (if selective), or constraints on escalation imposed by military rules of engagement. For insurgents, this parameter can be interpreted as the frequency of bombings, hit-and-run attacks, assaults, ambushes and assassinations, as well as larger-scale offensive operations.
- b) **Territorial control.** Selective violence is rarely as accurate as Figure 2a would suggest. A combatant's ability to correctly identify her rivals depends on the quality of her local intelligence, which in turn depends on the willingness of the local population to denounce supporters of the other side. As noted by Kalyvas (2008b, 407), such information is easier to obtain where the risk of retaliation, or counter-denunciation, is relatively low.¹¹ The quality of intelligence thus depends on the balance of territorial control in the conflict zone (θ_I, θ_G). Where territorial control is not absolute

¹¹The logic here is consistent with Kalyvas's insight that "individuals want to denounce [supporters of a rival faction] only where it is safe for them to do so. This is the case in areas of full control (where political actors do not need their information) but not in areas of low control (where they are likely to face retaliation)" (Kalyvas, 2008b, 407).

($0 < \theta_I < 1$ or $0 < \theta_G < 1$), as in Figure 2b, the “identification problem” asserts itself: some portion of overall punishment will befall combatants as intended ($\rho_I \theta_I, \rho_G \theta_G$), but the remainder will be erroneously inflicted on civilians ($\rho_I(1 - \theta_I), \rho_G(1 - \theta_G)$). Where combatant j exercises strong control ($\theta_j \rightarrow 1$), intelligence quality is relatively robust, enabling her to target opponents with higher precision ($\rho_j \theta_j \rightarrow \rho_j$) and avoid civilian casualties ($\rho_j(1 - \theta_j) \rightarrow 0$). Where the combatant is weak ($\theta_j \rightarrow 0$), her intelligence is poor and civilian deaths are harder to avoid ($\rho_j(1 - \theta_j) \rightarrow \rho_j$).¹²

- c) **Cooperation.** Civilians cooperate with insurgents at rate μ_I , and with the government at rate μ_G (Figure 2c). The relative magnitude of these rates depends on the strategy selected by civilians. If civilians choose to balance ($\beta = 1$), their cooperation with side A will be increasing in the rate of civilian victimization by side B. If they choose to bandwagon ($\beta = 0$), it will be increasing in the rate of victimization by side A. The two strategies present opposite incentives for combatants: balancing favors restraint, while bandwagoning favors escalation. If insurgents punish civilians at a low rate and the government punishes at a high rate, balancing civilians will be more inclined to cooperate with insurgents and bandwagoning civilians will be more inclined to cooperate with the government. Formally,

$$\begin{aligned}\mu_I &= \beta(1 - \theta_G)\rho_G + (1 - \beta)(1 - \theta_I)\rho_I \\ \mu_G &= \beta(1 - \theta_I)\rho_I + (1 - \beta)(1 - \theta_G)\rho_G\end{aligned}\tag{1}$$

As Figure 2c indicates, the overall scale of cooperation is proportional to the frequency of contacts between civilians and current active supporters ($\mu_I I, \mu_G G$). This proportionality assumes that physical contact between civilians and combatants is necessary for recruitment. It is not sufficient for an individual to unilaterally declare herself to be an active supporter, or for a combatant to simply add a civilian’s name to some roster. This definition excludes the “lone wolf” phenomenon, and assumes that combatants’ ability to recruit personnel is endogenous to existing levels of active support.

- d) **Immigration and Death.** In the absence of punishment, the civilian population is regulated by a simple immigration-death process. Civilians migrate into the conflict zone at rate k and are removed at a natural death rate u , which may be interpreted as losses due to disease, malnutrition, natural disasters, age and other exogenous factors

¹²For simplicity, I assume that territorial control is zero-sum ($\theta_I = 1 - \theta_G$ and $\theta_G = 1 - \theta_I$), such that one side can improve her intelligence assets only at the expense of the other. Consequently, as the use of force by one side becomes more accurate, use of force by the other side becomes more inaccurate.

that afflict civilians and combatants equally.¹³ The parameter k may be set as a constant rate, or a variable rate that balances exactly the death rates of all players $k_t = ((1 - \theta_I)\rho_I + (1 - \theta_G)\rho_G + u)C_t + (\theta_I\rho_I + u)G_t + (\theta_G\rho_G + u)I_t$.

Taken together, the dynamic process can be represented as a system of ordinary differential equations, in which the rate of change in popular support is a function of **cooperation**, **punishment** and **territorial control**:

$$\frac{dC}{dt} = k - (\mu_I I_t + \mu_G G_t + \rho_I(1 - \theta_I) + \rho_G(1 - \theta_G) + u)C_t \quad (2)$$

$$\frac{dI}{dt} = (\mu_I C_t - \rho_G \theta_G - u)I_t \quad (3)$$

$$\frac{dG}{dt} = (\mu_G C_t - \rho_I \theta_I - u)G_t \quad (4)$$

with μ_I, μ_G as defined in (1). Without loss of generality, let us assume that strategic choices are binary, such that $\rho_j = 1$ when combatant j plays H (high punishment) and $\rho_j = 0$ when she plays L (low). Similarly, let $\beta = 1$ when civilians play BL (balancing) and $\beta = 0$ when they play BW (bandwagoning).

Apart from the territorial control parameter and the endogenization of cooperation, the system in (2-4) resembles a traditional epidemiological model of infection biology, in which two parasite strains compete for the same host (Nowak and May, 1994, 2000; Nowak, 2006).

3 Civil War Outcomes

The outcome of the dynamic process in (2-4) depends on the relative size of the *basic reproductive ratio* for each combatant, defined as the number of new active supporters caused by the introduction of a single combatant into a population of neutral civilians. The basic reproductive ratios for insurgents and the government, respectively, are given by

$$R_I = \frac{\mu_I}{\rho_G \theta_G + u} \cdot \frac{k}{\rho_G(1 - \theta_G) + \rho_I(1 - \theta_I) + u} \quad (5)$$

$$R_G = \frac{\mu_G}{\rho_I \theta_I + u} \cdot \frac{k}{\rho_G(1 - \theta_G) + \rho_I(1 - \theta_I) + u} \quad (6)$$

with μ_I, μ_G as defined in (1). The first part of expression (5) represents the average number of new supporters recruited by an insurgent agent in her lifetime – the rate at which neutral

¹³The immigration-death process is frequently used in epidemiological models as a simple, though artificial, way to obtain a stable population (Nowak and May, 1994; May and Nowak, 1995).

civilians cooperate with a single insurgent (μ_I), scaled by the average lifetime of an insurgent supporter $((\rho_G\theta_G + u)^{-1})$. The second part represents the equilibrium abundance of neutral civilians – the rate of civilian immigration (k), scaled by the average lifetime of civilians $((\rho_G(1 - \theta_G) + \rho_I(1 - \theta_I) + u)^{-1})$. The interpretation is the same for the government in (6).

The basic reproductive ratio represents a critical threshold in epidemiology. If $R_I < 1$, fewer than one civilian will cooperate with each insurgent supporter and the insurgent population will converge to zero over time. If $R_I > 1$, the insurgent population will rise exponentially, peak and converge to a stable positive equilibrium. If $R_I < 1$ and $R_G < 1$, both combatant populations will dwindle and the system will converge to an equilibrium where everyone is a neutral civilian:

$$\text{Eq. 0:} \quad C_{eq} = \frac{k}{\rho_G(1 - \theta_G) + \rho_I(1 - \theta_I) + u} \quad I_{eq} = 0 \quad G_{eq} = 0 \quad (7)$$

If $R_I > 1, R_G > 1$ or both, one of three outcomes becomes possible: insurgent victory, government victory, or stalemate.

3.1 Conditions for insurgent victory

An insurgent victory is defined as an outcome where insurgents monopolize popular support and no government agents remain in the population, implying $I_{eq} > 0, G_{eq} = 0$, and $\pi_I(\cdot) = 1, \pi_G(\cdot) = 0$ as operationalized in Table 1. This equilibrium is represented by

$$\text{Eq. 1:} \quad C_{eq} = \frac{\rho_I\theta_I + u}{\mu_I} \quad I_{eq} = \frac{k\mu_I - (\rho_G(1 - \theta_G) + \rho_I(1 - \theta_I) + u)(\rho_G\theta_G + u)}{\mu_I(\rho_G\theta_G + u)} \quad G_{eq} = 0 \quad (8)$$

which is stable if and only if $R_I > R_G$ (see proof in appendix).

What strategy profiles are likely to produce this outcome? Where territorial control is evenly divided between the two combatants ($\theta_I = \theta_G = 1/2$), $R_I > R_G$ is true under the strategy profiles ($s_C = BL, s_I = L, s_G = H$) and ($s_C = BW, s_I = H, s_G = L$). In the first instance, an insurgent victory occurs when civilians balance against the more aggressive government. In the second, civilians bandwagon with the more aggressive insurgents. In each case, victory results from an asymmetric use of punishment ($\rho_I > \rho_G, \rho_G > \rho_I$).

An insurgent victory is also possible under symmetric punishment ($\rho_G = \rho_I > 0$), if one of the two sides enjoys an initial intelligence advantage. When insurgents have an advantage

in territorial control ($\theta_I > 1/2$), $R_I > R_G$ is true under ($s_C = BL, s_I = H, s_G = H$). When the government has an advantage ($\theta_G > 1/2$), $R_I > R_G$ is true under ($s_C = BW, s_I = H, s_G = H$). In the first instance, insurgents are able to more accurately identify and target government supporters, causing fewer civilian casualties than their enemy despite an equal rate of punishment. Victory occurs because balancing civilians flock to the side able that avoids collateral damage. In the second instance, insurgents are responsible for a greater share of civilian casualties, but win because bandwagoning civilians support the side most willing to hurt them.

3.2 Conditions for government victory

A government victory occurs when no insurgents remain in the population and the government monopolizes popular support ($I_{eq} = 0, G_{eq} > 0$, and $\pi_I(\cdot) = 0, \pi_G(\cdot) = 1$).

$$\begin{aligned} \text{Eq. 2:} \quad & C_{eq} = \frac{\rho_G \theta_G + u}{\mu_G} \quad I_{eq} = 0 \quad G_{eq} = \frac{k\mu_G - (\rho_I \theta_I + u)(\rho_G(1 - \theta_G) + \rho_I(1 - \theta_I) + u)}{\mu_G(\rho_I \theta_I + u)} \end{aligned} \quad (9)$$

This equilibrium is stable if and only if $R_I < R_G$ (see proof in appendix).

Assuming parity of territorial control ($\theta_I = \theta_G = 1/2$), $R_I < R_G$ holds under the strategy profiles ($s_C = BL, s_I = H, s_G = L$) and ($s_C = BW, s_I = L, s_G = H$). In the first instance, government victory occurs when civilians balance and the insurgents employ a higher rate of punishment. In the second, civilians bandwagon and the government uses more punishment than insurgents.

As before, local asymmetries produce two additional possibilities: ($s_C = BW, s_I = H, s_G = H$) if insurgents have an advantage in territorial control ($\theta_I > 1/2$), and ($s_C = BL, s_I = H, s_G = H$) if the government has the advantage ($\theta_G > 1/2$). In the first case, the government wins because civilians bandwagon with the more inaccurate side, and in the second civilians balance against it.

3.3 Conditions for stalemate

A stalemate occurs when popular support is evenly split between insurgents and the government at equilibrium ($I_{eq} = G_{eq}$, and $\pi_I(\cdot) = \pi_G(\cdot) = 1/2$). When $dI/dt = 0$ and $I > 0$, the equilibrium value of C is $(\rho_I \theta_I + u)/\mu_I$. When $dG/dt = 0$ and $G > 0$, the equilibrium value of C is $(\rho_G \theta_G + u)/\mu_G$. When both of these conditions are true simultaneously, we obtain

the following equilibrium:

$$\begin{aligned}
\text{Eq. 3: } C_{eq} &= \frac{\rho_I \theta_I + u}{\mu_I} = \frac{\rho_G \theta_G + u}{\mu_G} \\
I_{eq} &= \frac{k\mu_I - (\rho_G \theta_G + u)(\rho_G(1 - \theta_G) + \rho_I(1 - \theta_I) + u)}{\mu_I(\rho_G \theta_G + u)} \\
G_{eq} &= \frac{k\mu_G - (\rho_I \theta_I + u)(\rho_G(1 - \theta_G) + \rho_I(1 - \theta_I) + u)}{\mu_G(\rho_I \theta_I + u)} \\
I_{eq} &= G_{eq}
\end{aligned} \tag{10}$$

which is stable if and only if $R_I = R_G$.

Unlike the equilibria in (8-9), the stalemate outcome does not depend on civilian strategy and is completely determined by punishment choices and territorial control. At any level of incomplete control ($0 < \theta_I < 1$), $R_I = R_G$ if neither side punishes ($s_C = BL$ or BW , $s_I = L$, $s_G = L$). In areas of perfectly contested control ($\theta_I = \theta_G = 1/2$), $R_I = R_G$ if the rate of punishment is positive but symmetric ($s_C = BL$ or BW , $s_I = H$, $s_G = H$). When insurgents enjoy complete control ($\theta_I = 1$), enabling them to fully avoid civilian casualties, $R_I = R_G$ if ($s_C = BL$ or BW , $s_I = H$, $s_G = L$). Conversely, when the government enjoys complete control ($\theta_G = 1$), $R_I = R_G$ if ($s_C = BL$ or BW , $s_I = L$, $s_G = H$). In this sense, complete territorial control enables the hegemon to unilaterally punish without losing support, as long as the enemy doesn't escalate. Under such circumstances, the civilian population is unaffected by the violent interaction between combatants. It suffers no costs by way of collateral damage and remains indifferent as to which side it should support to maximize its security.

Figure 3 summarizes the game's payoffs under each strategy profile (s_C, s_I, s_G) , and under three distributions of territorial control: (a) divided control, $\theta_G = 1/2$, (b) incomplete government control, $1/2 < \theta_G < 1$, and (c) complete government control, $\theta_G = 1$.¹⁴ For civilians, the disutility associated with each strategy profile is increasing in the rate of punishment employed by the two sides, $\pi_C(s_C, s_I, s_G) = -(\rho_I(1 - \theta_I) + \rho_G(1 - \theta_G))$. Punishment is most costly if both combatants play H , least costly if both play L and intermediate if only one plays H . For combatants, payoffs are increasing in the equilibrium share of support, $\pi_I(s_C, s_I, s_G) = \frac{I_{eq}}{I_{eq} + G_{eq}}$ and $\pi_G(s_C, s_I, s_G) = \frac{G_{eq}}{I_{eq} + G_{eq}}$, with I_{eq}, G_{eq} as defined in (7-10).¹⁵

¹⁴Parameter values: $\beta = \{1 \text{ if } s_C = BL, 0 \text{ if } s_C = BW\}$; $\rho_I = \{1 \text{ if } s_I = H, 0 \text{ if } s_I = L\}$; $\rho_G = \{1 \text{ if } s_G = H, 0 \text{ if } s_G = L\}$; $k_t = (u + (1 - \theta_I)\rho_I + (1 - \theta_G)\rho_G)C_t + (u + \theta_I\rho_I)G_t + (u + \theta_G\rho_G)I_t$; $u = 1$.

¹⁵Insurgent and government victories ($\pi_I = 1, \pi_G = 0$ and $\pi_I = 0, \pi_G = 1$) occur under Eq. 1 and 2. A stalemate ($\pi_I = \pi_G = 1/2$) occurs under Eq. 0 or Eq. 3.

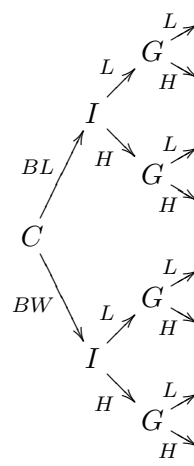
	(a) Divided $\theta_G = \frac{1}{2}$	(b) Incomplete $\frac{1}{2} < \theta_G < 1$	(c) Complete $\theta_G = 1$
			
$C \xrightarrow{BL} I \xrightarrow{L} G \xrightarrow{L}$	$(0, \frac{1}{2}, \frac{1}{2})$	$(0, \frac{1}{2}, \frac{1}{2})$	$(0, \frac{1}{2}, \frac{1}{2})$
$C \xrightarrow{BL} I \xrightarrow{H} G \xrightarrow{L}$	$(-\frac{1}{2}, 1, 0)$	$(-\theta_I, 1, 0)$	$(0, \frac{1}{2}, \frac{1}{2})$
$C \xrightarrow{BL} I \xrightarrow{H} G \xrightarrow{H}$	$(-\frac{1}{2}, 0, 1)$	$(-\theta_G, 0, 1)$	$(-1, 0, 1)$
$C \xrightarrow{BW} I \xrightarrow{L} G \xrightarrow{L}$	$(-1, \frac{1}{2}, \frac{1}{2})$	$(-1, 0, 1)$	$(-1, 0, 1)$
$C \xrightarrow{BW} I \xrightarrow{H} G \xrightarrow{L}$	$(0, \frac{1}{2}, \frac{1}{2})$	$(0, \frac{1}{2}, \frac{1}{2})$	$(0, \frac{1}{2}, \frac{1}{2})$
$C \xrightarrow{BW} I \xrightarrow{H} G \xrightarrow{H}$	$(-\frac{1}{2}, 0, 1)$	$(-\theta_I, 0, 1)$	$(0, \frac{1}{2}, \frac{1}{2})$
$C \xrightarrow{BW} I \xrightarrow{H} G \xrightarrow{H}$	$(-\frac{1}{2}, 1, 0)$	$(-\theta_G, 1, 0)$	$(-1, 1, 0)$
$C \xrightarrow{BW} I \xrightarrow{H} G \xrightarrow{H}$	$(-1, \frac{1}{2}, \frac{1}{2})$	$(-1, 1, 0)$	$(-1, 1, 0)$

Figure 3: **Payoffs:** (a) parity of territorial control, $\theta_G = 1/2$, (b) incomplete government control, $1/2 < \theta_G < 1$, (c) complete government control, $\theta_G = 1$. Numbers in parentheses represent payoffs in the order (π_C, π_I, π_G) .

The game tree shows how civilian strategy shapes incentives for punishment. If civilians balance, whichever side can minimize its share of civilian casualties will win the game. Under bandwagoning, whichever side can terrorize civilians the most will win the game. Yet if combatant victory, defeat and stalemate can occur under either of the two civilian strategies, two questions arise. First, is it ever preferable for civilians to bandwagon rather than balance? Second, if – as one might expect – civilians prefer a strategy that rewards restraint, why would we ever observe high rates of punishment in equilibrium?

4 When Does Punishment Occur?

The following section explores the equilibria in (8-10) as outcomes of a strategic interaction between civilians and combatants. It is shown that the strategy profile $s_C = BL, s_I = L, s_G = L$ represents a subgame perfect equilibrium. Under perfect information, civilians' best response survival strategy is to balance, while combatants are both best off using a low level of punishment. This equilibrium, however, can become unstable if insurgent and government supporters are not aware of the civilians' strategy choice. Under imperfect information, the (BL, L, L) equilibrium breaks down and high rates of punishment occur. This punishment is likely to be two-sided where territorial control is evenly split, and one-

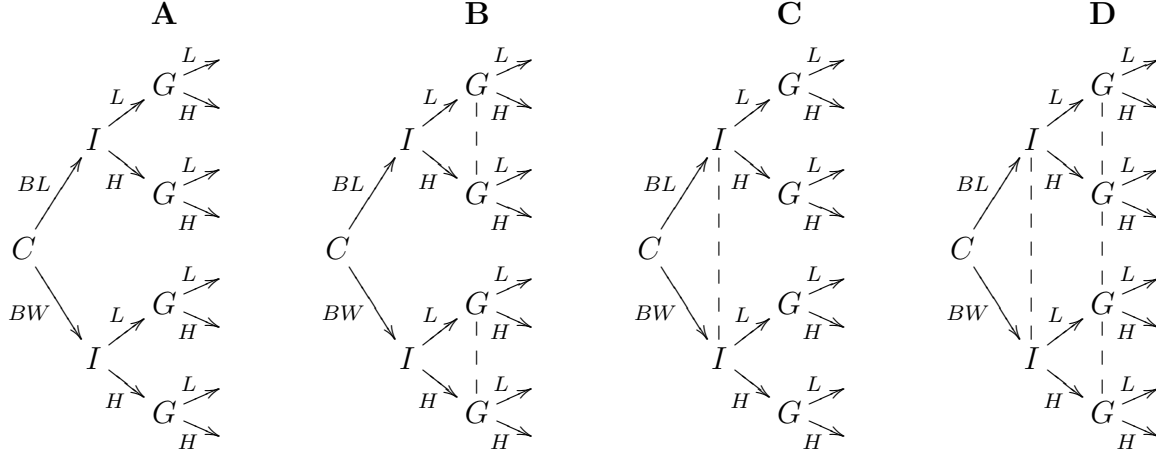


Figure 4: **Informational Assumptions.** Dotted lines represent information sets available to each player. If a dotted line connects two or more decision nodes, the corresponding player is aware that the preceding player has moved, but does not know which node within the information set has been reached as a result of that player's move.

sided where the balance of control favors one of the two combatants.

We now examine best response strategies under four sets of informational assumptions:

- A. Complete and perfect information. Civilians choose to balance or bandwagon, while insurgent and government supporters choose optimal levels of force. The preferences and choices of each player are visible to all other players (Figure 4a).
- B. Combatant choices unknown. Civilians' choice is observed by the insurgents and the government, but neither combatant can observe what the other does (Figure 4b).
- C. Civilian choice unknown. Combatants observe each other's moves, but not the civilians' choice to balance or bandwagon (Figure 4c).
- D. All choices unknown. Civilians, insurgents and the government do not observe any of the choices taken by other players (Figure 4d).

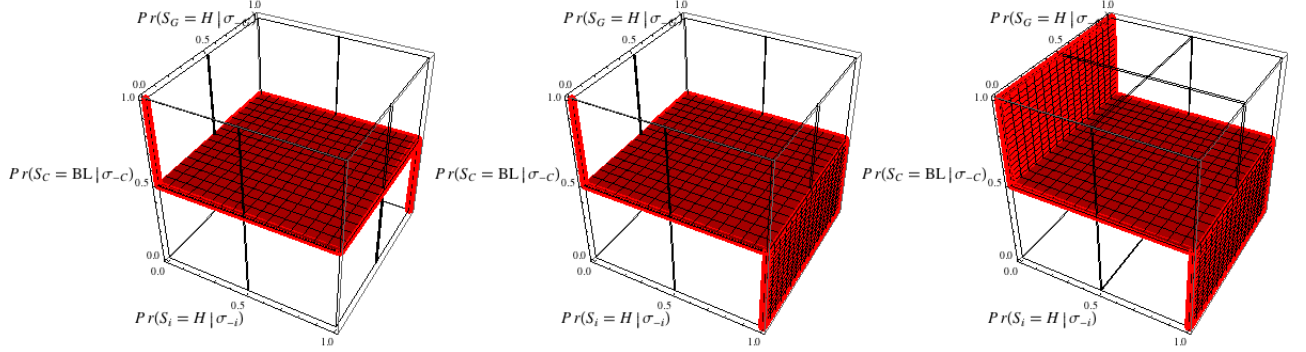
4.1 Divided territorial control

We begin with the case where neither side enjoys an advantage in territorial control ($\theta_G = 1/2$). Here, when each actor is given perfect information about the decisions of previous players (condition A), the strategy profile (BL, L, L) constitutes a subgame perfect equilibrium. The civilians have the upper hand: knowing that insurgents are likely to play L

if civilians balance and H if they bandwagon, civilians will prefer the former equilibrium because it minimizes the costs of being punished. Knowing that civilians have selected a balancing strategy, L is unconditionally best-performing in the resulting subgame. Insurgents understand that if they opt for a low level of violence, the government will do the same, and if they opt for a high killing rate, the government will choose restraint. Since insurgent payoffs are maximized when the combatants play LL rather than HL , insurgents will choose the former, resulting in the profile (BL, L, L) . No single player or coalition of players can accomplish a Pareto-improvement by deviating from this equilibrium.

The same result holds when the players are given more limited information about each other's choices (condition B). Here, civilians' choices are observed, but the two combatants' moves are not. Even so, the dominant strategy profile remains $s_C = BL$, $s_I = L$, $s_G = L$. In a balancing scenario, each combatant will be better off playing L irrespective of the other's choice. In a bandwagoning scenario, each will be better off playing H . Civilians anticipate that some Nash equilibrium will be played in each of the subgames in the second stage. Because they prefer an LL outcome to HH , civilians will choose to balance.

The uniqueness of the (BL, L, L) equilibrium falls apart when combatants are not given information about the civilians' choices (condition C). Here, insurgents' moves are observed by the government, but neither combatant is aware of which subgame they are playing – balancing or bandwagoning. The solution becomes a mixed strategy equilibrium in which one of four outcomes is possible: (BL, L, L) , (BL, H, H) , (BW, L, L) and (BW, H, H) , where (BL, L, L) Pareto-dominates the others. Insurgents understand that if they play L or H , the government will look at the potential outcomes in both subgames and will play whichever strategy maximizes her minimum payoff. The minimum government payoff possible under an LL profile is $1/2$, compared to 0 under LH ; the minimum payoff under an HL scenario is 0 , compared to $1/2$ under HH . As a result, insurgents expect the government to match their rates of killing. Insurgents further understand that – while their payoffs are identical under LL and HH – the stability of each equilibrium depends on the civilian's choice. Instead of choosing one of the two pure strategies, insurgents will choose some lottery among them, knowing that the government, as the final mover, will simply copy this choice. Civilians, unsure of whether the combatants will play LL or HH , will choose whichever strategy maximizes their lowest potential payoffs from the two scenarios. Under both balancing and bandwagoning, however, this minimum payoff is -1 . To cope with this indifference and the uncertainty surrounding the insurgents' decisions, civilians will seek to optimize their chances of survival by randomizing their use of balancing and bandwagoning. This result is equiva-



(a) **Divided control** ($\theta_G = 1/2$) (b) **Incomplete control** ($\theta_G > 1/2$) (c) **Complete control** ($\theta_G = 1$)

Figure 5: **Mixed strategy equilibria.** Red areas represent the intersection of players' best response mappings.

lent to what would obtain if the order of moves were to be reversed, with civilians moving last.

This example illustrates that punishment can emerge when combatants are not aware of civilians' choices. Although the equilibria (BL, L, L) and (BW, H, H) do not present any opportunities for profitable unilateral or coalition deviations, the two other equilibria are inherently unstable. In a (BL, H, H) equilibrium, both combatants have an incentive to unilaterally switch to an L strategy. In a (BW, L, L) equilibrium, both would be better off unilaterally switching to H . Unless combatants can obtain information on civilian decision-making as they did in sets A and B, large-scale punishment can be difficult to prevent.¹⁶

Restraint becomes even more difficult to sustain when no player has information on the choices of any other (condition D), and strategic decisions are either unknown or unobservable. Let $\sigma(s)$ be the joint probability associated with strategy profile $s = (s_C, s_I, s_G)$ when civilians, insurgents and the government play mixed strategies. Let $Pr(s_j | \sigma_{-j})$ be

¹⁶A natural objection to this result is that real-life actors do not play mixed strategies. Civilians and insurgents are unlikely to leave their fate to a coin toss, and will rely instead on their prior beliefs about each others' behavior. While such objections are perfectly valid and justified, I assumed flat priors for two reasons. First, there is no shortage of potential mechanisms for the formation of actors' beliefs – “focal points” (Schelling 1960), fear, salience (Sandler and Sargent 1995) and Bayesian updating are four among many (Medina 2010, 44-51). Informative priors must come from somewhere, and this origin must be specified and defended against its alternatives. This type of exercise, while useful, is ultimately outside the scope of the current paper. Second, as with any model being introduced for the first time, core predictions are most easily presented and interpreted without potentially confounding assumptions about players' culture, institutions, history and traditions.

the probability that player j chooses strategy s_j , conditional of the mixed strategies of all other players. The full set of mixed strategy equilibria is shown graphically in Figure 5a, as an intersection of players' best response mappings (Ungureanu and Botnari, 2005). When $Pr(s_C = BL|\sigma_{-C}) > .5$, the combatants will play LL with probability 1. When $Pr(s_C = BL|\sigma_{-C}) < .5$, the combatants will play HH with probability 1. When $Pr(s_C = BL|\sigma_{-C}) = .5$, the combatants will choose some lottery of L and H . Asymmetric violence (LH or HL) occurs in equilibrium only if civilians are equally likely to balance or bandwagon. In every other case, we should expect to see two-sided restraint (LL) or two-sided punishment (HH).

4.2 Incomplete government control

The results shown thus far assume that neither combatant benefits from an advantage in territorial control, and the ability to collect intelligence and correctly distinguish one's opponents from civilians is identical for the two sides ($\theta_I = \theta_G = 1/2$). What sorts of interactions might we expect where the balanced of power is more uneven? Because the combatants' subgame is symmetric, I focus here on the case of incomplete government control, with the understanding that results generalize to insurgent advantage as well.

A higher level of government control yields an intelligence advantage, reducing the insurgents' ability to accurately identify opponents ($\theta_I < 1/2$) and increasing the government's ability to identify insurgent collaborators ($\theta_G > 1/2$). Even if they exactly matched the government's level of violence (HH), insurgents would still be responsible for a higher share of civilian casualties, and government forces could successfully eliminate a greater share of insurgent supporters. Because civilian costs from insurgent punishment are relatively high, balancing civilians will cooperate with the government. Bandwagoning civilians will cooperate with insurgents.

Returning to the four sets of informational constraints discussed above, intelligence asymmetry does not change the outcome of the game so long as civilian choice is known (conditions A and B). In the balancing subgame, LL is the Nash equilibrium. The bandwagoning subgame has two equilibria: HL and HH . Civilian costs are lowest under LL , leading to the unique equilibrium (BL, L, L) .

When civilian choice is unknown (conditions C and D), additional outcomes become possible. Unlike in the case of intelligence parity, insurgents cannot expect the government to copy their punishment strategy. The minimum government payoff possible under an LL profile is $1/2$, which is preferred to a minimum of 0 under LH . However, the government's

minimum payoff – like that of the insurgents – is 0 under both HL and HH . Insurgents understand that the government is likely to respond to L by playing L , and to H by playing some lottery of L and H . Risk-averse insurgents will prefer a worst-case scenario of stalemate to a worst-case scenario of defeat, leading to a combatant strategy profile of LL . While civilian payoffs from (BL, L, L) and (BW, L, L) are identical, the (BW, L, L) equilibrium is far more unstable – both insurgents and the government can obtain a Pareto improvement by unilaterally escalating punishment. Indeed, insurgents can secure victory through escalation irrespective of whether the government also deviates from L .

The resulting set of mixed strategy equilibria (Figure 5b) is more expansive than that under parity (Figure 5a), though potentially less costly to civilians. As before, combatants will both play L if the probability of balancing is better than even ($Pr(S_C = BL|\sigma_{-C}) > .5$) and some lottery of L and H if $Pr(S_C = BL|\sigma_{-C}) = .5$. If civilians are more likely to bandwagon ($Pr(S_C = BL|\sigma_{-C}) < .5$), mutual punishment (HH) is no longer assured. If civilians bandwagon with the side that inflicts the most costs, then the insurgents – if only by virtue of their local disadvantage – are always better off escalating. The government, meanwhile, is indifferent to the payoffs under (BW, H, L) and (BW, H, H) , since they lose support in either case. As long as $Pr(S_G = H|\sigma_{-I}) < 1$, however, expected costs to civilians under bandwagoning ($Pr(S_C = BL|\sigma_{-C}) < .5$) are lower in areas of incomplete control than in areas of parity: an equilibrium in which only one side always punishes is less costly than one in which both combatants always punish. In this limited sense, bandwagoning civilians living in areas of partial control are better off than those in areas of divided control.

Initial advantages in territorial control, as these results show, do not translate easily into victory for the side that possesses them. Such an outcome can only occur if civilians balance and both sides punish at the same rate – a strategy profile that cannot be maintained in equilibrium. Oddly enough, the disadvantaged side can benefit more from its weakness than the advantaged side can from its strength. As Figure 5b shows, bandwagoning encourages unconditional punishment by the weaker side. To terrorize the population into lending its support, the disadvantaged insurgents need only to match or exceed the government’s level of violence.

4.3 Complete government control

Do these results generalize to non-contested areas, where one of the combatants exercises a monopoly on territorial control? Under complete control ($\theta_G = 1$), the government is able to perfectly monitor the population and correctly identify insurgent supporters. Because

the identity of insurgents is public knowledge, the government is unrestrained in the level of violence it can use against them. The number of civilian casualties is zero under L and H , enabling the government to unleash a wave of arrests or executions without fear of alienating a balancing civilian population. Because payoffs are unaffected by the government's use of force, the initiative now lies with insurgents, who have no access to the population and can only use highly inaccurate forms of violence to obtain coercive leverage.

When civilian choice is known (conditions A and B), two outcomes are possible: (BL, L, L) and (BL, L, H) . In the balancing subgame, L is the insurgents' unconditionally best-performing strategy, but the government is indifferent between L and H . In the bandwagoning subgame, insurgents are always best off playing H , but the government is again indifferent. Civilians expect costs of 0 under BL and -1 under BW , and choose the former.

When civilian choice is unknown (conditions C and D), the game has four potential outcomes – (BL, L, L) , (BL, L, H) , (BW, H, L) and (BW, H, H) – and a broad space of mixed strategy equilibria (Figure 5c). If the probability of balancing is better than even ($Pr(S_C = BL|\sigma_{-C}) > .5$), insurgents always play L and the government plays a lottery of L and H . If $Pr(S_C = BL|\sigma_{-C}) < .5$, the insurgents play H and the government plays L or H . If $Pr(S_C = BL|\sigma_{-C}) = .5$, both combatants play some lottery of L and H .

While local asymmetries in territorial control add a layer of complexity to the players' strategic calculus, they do not fundamentally alter the logic of the game: combatants are likely to play L when civilians balance and H when civilians bandwagon. The key distinction is whether the resulting violence is unilateral or two-sided. Under incomplete control, civilian balancing deters punishment by both sides and bandwagoning encourages escalation, especially for the weaker side. Under complete control, bandwagoning has the same effect, but balancing deters only the weaker combatant: as long as the government's use of violence is perfectly selective, inflicting no costs on neutral civilians, the hegemon neither loses nor gains support by cracking down on her opponents.

5 The evolution of punishment

The solutions examined above assume that the three players act in a unitary fashion, are aware of the payoffs associated with each strategy profile, and will select whichever strategy optimizes these payoffs given the choices available to other players. In practice, however, strategy choice is often a matter of trial and error: based on a prior history of strategic interactions, players will adopt well-performing strategies and abandon poor-performing ones.

Over time, this learning process should converge to a steady state, where dominated strategies will have mostly disappeared from the players' repertoire. In place of a Nash Equilibrium from classical game theory, such evolutionary games turn on the diffusion of best practices.¹⁷

To examine the likelihood of punishment in an evolutionary context, I develop an agent-based model in which 50 actors of each type are randomly grouped in sets of three (C, I, G), and play the game as described in condition D in Section 4.1. The initial strategy for each of the 150 players is decided by a fair coin toss (BL or BW for the civilians, L or H for the combatants). For each of t generations, the randomly grouped agents play the game against each other for 100 rounds, and receive payoffs π_C, π_I, π_G associated with the triad's strategy profile. The players then evaluate their strategies: the player with the lowest payoff in each of the groups C, I, G abandons her strategy and adopts the strategy of the best-performing player of her type.¹⁸ Of the players who switch strategies, some proportion p choose a random "mutant" strategy instead of the incumbent, best-performing one. The players are then randomly re-grouped into new triads. This cycle is repeated for 10,000 generations, over the course of which successful strategies increase in frequency at the expense of the less successful. Unless "mutant" strategies outperform incumbent ones, the population should converge to an evolutionarily stable state, where the proportion of agents playing pure strategy s_j can be interpreted as a mixed strategy equilibrium.¹⁹

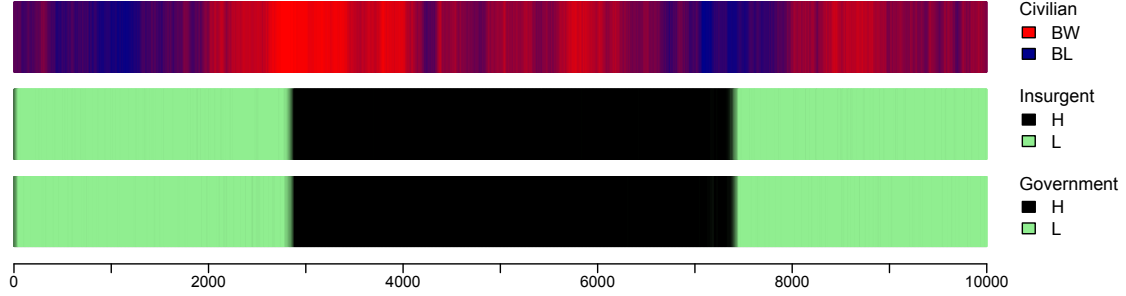
Figure 6 shows the population's evolution over 10,000 generations under three scenarios: (a) divided territorial control, (b) incomplete government control, and (c) complete government control. In each case, the mutation parameter is set at $p = .05$. The horizontal axis displays the generation number, while the color ramp indicates the proportion of agents in each group (C, I, G) playing each of the pure strategies. Solid blue lines indicate that 100% of the civilian population is using a balancing strategy, while solid red indicates that 100% is bandwagoning. Purple lines indicate a mixed, or polymorphic population. For combatants, green lines indicate that 100% of the actors are using a low level of force, and black lines indicate that 100% are using a high level of force. At the outset of the simulation, each group is split evenly between balancers and bandwagoneers, or high and low punishers.

The results of the agent-based model show that a peaceful equilibrium can be difficult to maintain due to the fickleness of civilians. In the divided control scenario (Figure 6a), the

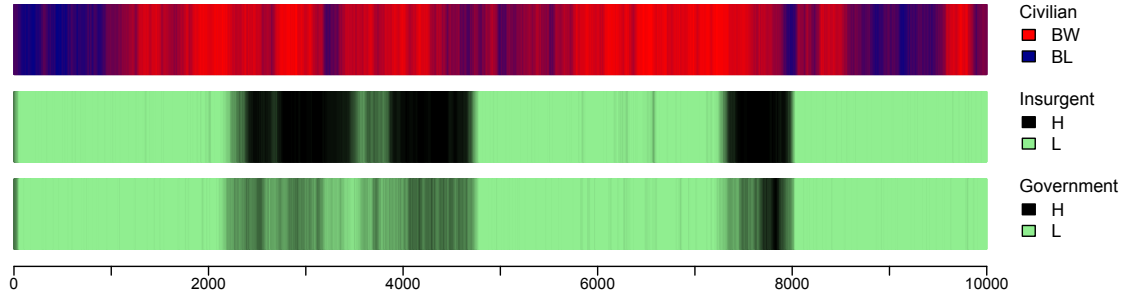
¹⁷See Maynard Smith (1982); Gintis (2009).

¹⁸In the case of ties, the worst- and best-performing players are selected at random, proportional to their payoffs.

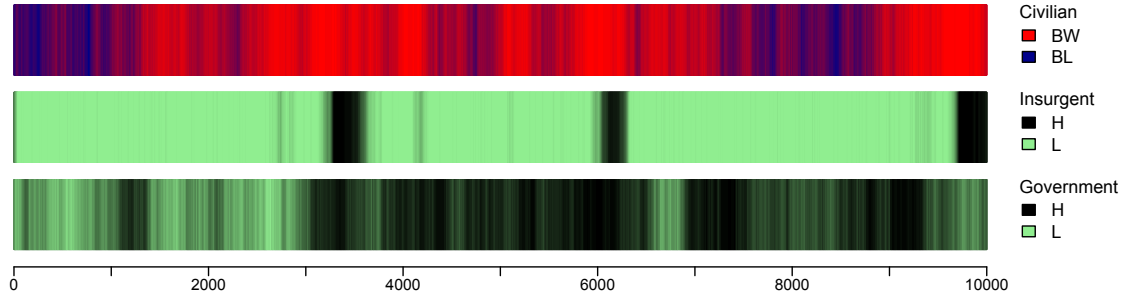
¹⁹Much like a Nash Equilibrium presupposes that no player can benefit for unilateral deviation, an Evolutionarily Stable Population State (ESPS) presupposes that a population cannot be "invaded" by players with mutant strategies.



(a) Divided control ($\theta_G = 1/2$)



(b) Incomplete government control ($\theta_G = 3/4$)



(c) Complete government control ($\theta_G = 1$)

Figure 6: **Evolutionary agent-based model.** Horizontal axis represents time (in generations), color shades represent proportion of population playing each pure strategy.

combatants' subgame converges to the state (L, L) after fewer than 100 generations, while over half of the civilian population plays BL . This state is akin to the upper-right hand corner of the mixed strategy equilibrium space in Figure 5a – if the probability of balancing is greater than even, both combatants will choose a low rate of punishment. Civilians, however, prove highly vulnerable to perturbations from mutation and have difficulty coalescing around a single strategy. As shown before in Figure 3, civilian payoffs are identical under the outcomes (BL, L, L) and (BW, L, L) . As a result, civilians playing BW perform about as well as those playing BL on average, provided that the combatants do not escalate. Due to this indifference, balancing and bandwagoning strategies coexist in the population. If the share of bandwagoning civilians becomes sufficiently high – as it does here after about 2500 generations – combatants begin to realize the benefits of deviation from L . As mutant combatants playing H have more chances to interact with bandwagoning civilians, they outperform combatants playing the incumbent L strategy. By generation 3000, other combatants catch on to the benefits of escalation, and the system shifts from the low-violence state (BL, L, L) to the mass killing state (BW, H, H) . After this happens, civilians gradually begin to re-adopt a balancing strategy to maximize their survival. The HH equilibrium, however, can endure for thousands of generations before peace is restored. Rather than settling into a single enduring state, the dynamics of the system are characterized by multiple equilibria: periods of fragile peace interrupted by long spells of two-sided violence.

Similar dynamics are seen in the case of incomplete government control (Figure 6b), but the distribution of violence is more one-sided. The combatants begin by adopting an LL profile while a majority of civilians plays BL . As the share of civilians playing BW increases, mutant strategies of H begin to outperform the incumbent L . As we would expect from the mixed strategy equilibria in Figure 5b, the proliferation of H strategies is noticeably greater among the disadvantaged side. In the violent spells that begin at approximately the 2200th and 7200th generations, nearly all insurgents adopt a high rate of violence, while only between a quarter and half of government agents do the same. When the balance of territorial control partially favors one side, the state of the system alternates between periods of peace and mostly one-sided violence, perpetrated by the weaker combatant. Because it is no longer two-sided, violence in this region is noticeably less costly to civilians than in areas of divided control: average civilian payoffs during the first 10,000 generations were -0.244 when $\theta_G = 3/4$, compared to -0.458 when $\theta_G = 1/2$.

In the case of complete government control (Figure 6c), insurgents again take advantage of civilian drifts into bandwagoning. The government, however, is able to follow a wholly in-

dependent strategic path. Regardless of civilian strategy, government agents who play H do no better and no worse than those who play L , and the evolutionary process does not result in strategic convergence. While the insurgent population is relatively homogenous in playing L unless a sufficiently high number of civilians bandwagon, highly-violent government agents continuously coexist with peaceful ones. Because the government’s use of punishment is highly accurate, however, civilians suffer the least in areas of complete control: average civilian payoffs were -0.099 when $\theta_G = 1$, less than half of what they were under incomplete control. Although these costs could be completely avoided if civilians never bandwagoned – as would obtain under perfect information – the consolidation of sovereignty does make bandwagoning less risky, protecting civilians from their own worst instincts.

6 Illustrative example: Soviet counterinsurgency

Are the model’s main predictions consistent with the empirical record? The Soviet campaign against the Organization of Ukrainian Nationalists (OUN) and its military arm, the Ukrainian Insurgent Army (UPA) offers a fitting opportunity to compare the model’s analytical and computational results against stylized historical facts. The nationalist insurgency – which primarily gripped the eight western regions (*oblasts*) of Ukraine during the late stages of World War II and the subsequent period of post-war reconstruction (1943-50) – was the longest and most destructive domestic conflict encountered by the Soviet Union since its founding in 1922.²⁰ Over its course, both the insurgents and the Soviet People’s Commissariat for Internal Affairs (NKVD) placed a strategic emphasis on the punishment of suspected enemy collaborators.²¹ As the balance of territorial control shifted from parity to partial and then near-complete Soviet control, opportunities and motivations for punishment changed in telling ways.

The following analysis relies on event data assembled from declassified incident reports, war diaries, detainee interrogation transcripts and after-action reports from the Main Directorate of the NKVD.²² These internal-use documents offer a rare glimpse at the real-time

²⁰The conflict region includes Chernovitskaya, Drobychskaya, Lvovskaya, Rovenskaya, Stanislavskaya, Tarnopol’skaya, Volynskaya and (after 1946) Zakarpatskaya oblasts.

²¹While the OUN was a Ukrainian nationalist organization, the chosen case does not violate the model’s assumptions about cooperation: with the notable exception of Poles – who had been largely removed from the region through earlier campaigns of ethnic cleansing – local civilians generally had the option of cooperating with either combatant. Indeed, the ranks of the OUN included a significant number of Red Army deserters and ex-POWs from other parts of the USSR (Vladimirtsev and Kokurin, 2008, 425).

²²Key archival data sources include GARF R-9401, Op. 1-2; GARF, F. R-9478, Op. 1; GARF, F. R-9479, Op. 1; RGVA, F. 38650, Op. 1; TsDAGOU, F. 1, Op. 23

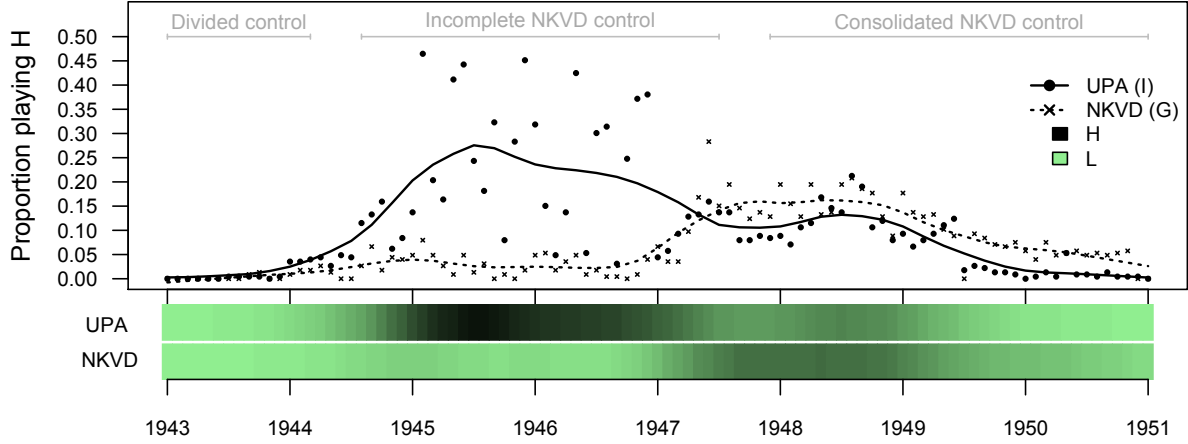


Figure 7: Time series, Combatant strategy choices in West Ukraine, 1943-1950.

information available to Soviet and insurgent commanders over the course of the conflict. The subset used here includes 7,132 conflict events between 1943 and 1950 that meet Kalyvas (2006)’s definition of selective violence,²³ each with micro-level information on locations, dates, casualties and tactics. For the following analysis, these events were aggregated to time-series cross-sectional data at the level of a district (*rayon*)-month.²⁴

The dynamics of violence are shown in Figure 7, where the x -axis represents time and the y -axis is the proportion of rayons where either the insurgents (UPA, solid line) or the government (NKVD, dotted line) employed a high rate of punishment, defined as the incidence of at least one episode of selective violence per month. The points represent unfiltered monthly observations (proportion of districts where each combatant plays H) and the smoothed lines represent the trend component of the time series, extracted with loess regression as part of a seasonal-trend decomposition (Cleveland et al., 1990). Below the line plot is an alternate view of the time trend, with the same symbology as in Figure 6. Light green shades indicate that combatants in most rayons play L and darker shades indicate that they play H .

The Ukrainian case can be separated into three phases: (1) a period of parity between the

²³Events were coded using an actor-action-target framework. For an event of *selective insurgent violence*, actors included OUN-SB (security service) and UPA; tactics included firefights, encirclements, assassinations, raids, ambushes, defensive battles, kidnappings, and acts of terrorism; targets included security forces, party activists and civilians. For *selective government violence*, actors included the internal troops of the NKVD/MGB/MVD, Soviet partisans, and paramilitaries (“destruction battalions”); tactics included firefights, encirclements, assassinations, raids, ambushes, weapons seizures, defensive battles and deportations; targets included suspected insurgents, their families and civilians. This definition excludes incidents of *indiscriminate violence*, which do not require the provision of information by civilians (e.g. cordon-and-search operations, search-and-destroy missions, artillery and air strikes, large-scale positional or maneuver warfare).

²⁴Each cross-section is based on a universal sample of 225 rayons from the eight West Ukrainian oblasts (Presidium of Supreme Soviet of USSR, Information-Statistical Division, 1946).

OUN and Soviet partisans in 1943, (2) a period of dominant, but incomplete Soviet control in 1944-1947, and (3) a period of consolidated Soviet control in 1948-1950. The epidemic model predicts symmetric levels of selective violence under divided control, and asymmetric levels of violence under incomplete and total control, with the disadvantaged side more likely to escalate in the first case and the dominant side more likely to escalate in the second. The dynamics of the Ukrainian conflict are largely consistent with these expectations.

The most common strategy profile during the period of divided control in 1943 was ($s_C = BL, s_I = L, s_G = L$). Although pre-war Soviet institutions were never robust in the borderland regions of Western Ukraine – which were annexed from Poland after the Nazi-Soviet Pact of 1939 – the local party apparatus was completely dismantled under German occupation. The main local agents of the government during this period were Soviet partisans, who launched their first raids in the region during the autumn of 1942. The OUN, which sought to establish an independent Ukrainian state, saw the partisans as a more dangerous rival than the German security forces, particularly as the local population grew increasingly hostile to the Nazis and were eager to support any force that would militarily challenge them (Statiev, 2008). In response to partisan raids, nationalist forces loyal to Stepan Bandera (OUN-B) organized their own armed militia (UPA) in late 1942, and focused the bulk of their subsequent military activity against Soviet agents.

As the prospect of German defeat became more apparent in 1943, the local population gradually became polarized between pro- and anti-Soviet elements (Statiev, 2010, 78-79). Although nationalist guerillas were initially more numerous than the partisans, they were not as well organized and faced a public relations problem due to perceptions of collaboration with the Germans. The UPA was indeed effective at frustrating partisan operations against the Wehrmacht, but also faced a war on multiple fronts against internal political enemies, ethnic Poles, Soviet collaborators and – for six months in 1943, largely in response to public pressure – against occupying German authorities (Statiev, 2008). Given the population's response to repressive German occupation policies (Dallin, 1981), an expectation of civilian balancing prevailed. Not until the Red Army's reoccupation of the borderlands in February-August 1944 did the insurgents turn their full attention to the large-scale punishment of suspected Soviet supporters.

The dominant strategy profile under incomplete Soviet control in 1944-1947 was (BW, H, L). As the Soviets reasserted their presence in the borderlands, establishing local party councils (*sel'sovety*) and drafting military-age men into the Red Army, the government at first abstained from the use of selective violence and sought to fight the UPA with conventional

means like positional battles, search-and-destroy missions and cordon-and-search operations (Vladimirtsev and Kokurin, 2008, 136). This approach – which did not rely on intelligence from local civilians – brought significant early successes, decimating the larger formations of the UPA.²⁵ Suffering catastrophic losses and large-scale defections to the Soviets, the OUN-B re-organized the UPA into small, mobile units suitable for guerilla warfare and began a campaign of terror and intimidation against suspected Soviet agents. Groups selected for insurgent punishment included “Komsomol members, Red Army officers, policemen, [...] those who evade service in UPA, along with their families,” “collectivization activists,” agricultural specialists dispatched from East Ukraine, peasants who conceded to Soviet grain requisitions or failed to deliver food supplies to the UPA, and civilians who paid government duties, voted in local elections or were even slightly suspected of treason (Statiev 2010, 124; Dyukov et al. 2009, 16-17). While this violence was selective by intent, the overwhelming majority (74 percent) of the insurgent attacks in Figure 7 were directed at civilians.

If the OUN selected this approach on the expectation of civilian bandwagoning, this assumption at first proved justified. As one participant wrote, “there is no point in doing political work in areas ... where [the OUN] perpetrates such violence” (Statiev, 2010, 129). The proportion of insurgents voluntarily surrendering to Soviet authorities (as opposed to those killed in action or captured) declined from 32 per cent to 16 per cent between 1945 and 1946.²⁶ In an environment of constant terror, the Soviets had great difficulty raising local cadres. Some rayon-level administrations operated with less than half of essential personnel, with no courts, prosecutors and an understaffed district NKVD office (Burds, 1997, 113-114). The resulting inability to collect reliable intelligence nullified Soviet advantages in firepower and excluded the types of selective violence that could eradicate the OUN’s network of small cells. Instead, the NKVD continued a strategic emphasis on massive operations that did not depend on the flow of actionable intelligence from local informants. As of mid-1946, the vast majority of insurgent actions went uninvestigated and three-fourths of all NKVD operations resulted in no contact with the enemy.²⁷

Around the same time, however, civilian strategy began to shift. As noted in an NKVD situation report from March 1946, “The population significantly altered its attitude toward the OUN [...] We have recorded a number of cases where local residents offered direct assistance to [our] forces in locating and liquidating the bandits [and] refused to deliver food to the bandits. Many bandits, witnessing the change [...] and fearing being surrendered to

²⁵TsDAGOU, F. 1, Op. 23, D. 2967, L. 25.

²⁶TsDAGOU, F. 1, Op. 23, D. 2967, L. 25. 171.

²⁷TsDAGOU, F. 1, Op. 23, D. 2966.

organs of Soviet power, relocate to other rayons and villages where no one knows them.”²⁸ This shift from bandwagoning to balancing enabled the NKVD to greatly expand its informant network and, after an internal review in January 1947, adopt a strategy that depended on more selective forms of violence (Vladimirtsev and Kokurin, 2008, 369).

As the Soviets consolidated control in 1947-1950, the strategy profile gradually converged to (BL, L, H) . Having conducted an initial census of families of suspected and known OUN members as early as March 1944, the NKVD and its successor, the MVD, greatly escalated the practice of forcible deportations in 1947.²⁹ Whereas deportations elsewhere in the Soviet Union, particularly in the North Caucasus, were notorious for indiscriminately uprooting hundreds of thousands of civilians on the sole principle of nationality, the deportations in the borderlands were of a more selective character. They were smaller, more frequent, and mostly limited to guerrilla relatives and active supporters. Whereas the national deportations were largely unconditional applications of brute force, the new deportations were used as instruments of compellence, often avoidable if wayward relatives surrendered to authorities (Statiev, 2005). While earlier waves of guerilla deportations were generally limited to 500 families or less, these policies would now assume a massive scale: the first wave of large-scale deportations in autumn 1947 relocated 26,644 guerilla families, or 76,192 individuals.³⁰

As the scale of Soviet punishment rose to unprecedented heights, authorities were careful not to provoke a new backlash among the population. Excesses and cases of civilian casualties were promptly blamed on inept local officials and the nationalist underground (Burds, 1997, 128-129). In some cases, security forces sought to exploit the population’s balancing tendencies by conducting raids on villages while dressed as UPA insurgents. Yet by 1949 even this practice would be abandoned as “blatantly provocative and imprudent.”³¹ Meanwhile, the MVD expanded its efforts to ‘Ukrainize’ the conflict by recruiting local cadres for administrative positions, paramilitary ‘extermination battalions’ and self-defense forces. The authorities were able to greatly expand their network of informants on the local level, recruiting numerous UPA defectors and captured insurgents. These improved intelligence assets helped the Soviets overcome identification problems, avoid excessive casualties among civilians, while inflicting heavy losses on OUN leadership, most notably the UPA’s supreme commander Roman Shukhevych, who was killed in an MVD ambush in March 1950.

As the Soviets became able to target insurgents with increased accuracy, the OUN found

²⁸RGVA, F. 38650, Op. 1, D. 145, L. 121-6.

²⁹GARF, F. R-9401, Op. 2, D. 64, L. 216-218.

³⁰GARF, F. R-9401, Op. 2, D. 199, L. 232-236.

³¹TsDAGOU, F. 1, Op. 16, D. 68, L. 9-17.

itself increasingly isolated. By 1949 the nationalists' military capabilities had been greatly diminished. Desertion and suicide rates were high, and internal revolts against commanders became frequent (Koval, 2003, 73). The principal targets of insurgent punishment remained civilians, who were often raided for no discernible purpose other than the insurgents' own subsistence. At the end of 1949, the OUN supreme leadership ordered a general demobilization of UPA and a halt to all guerilla activity (Tys-Krokhmaluk, 1972, 310). While the MVD continued its policy of targeted arrests and assassinations against remaining pockets of die-hard nationalists, the West Ukrainian insurgency had effectively been defeated.

7 Conclusion

The preceding analysis explored the logic behind two responses to civil war violence: balancing against the side that inflicts the most costs, and bandwagoning with it. Using an epidemic model of popular support dynamics and solution concepts from game theory, I showed that – in a world of perfect information – security-minded civilians are always better off balancing, and neither the insurgents nor the government has an incentive to escalate the use of force. Bandwagoning, which encourages escalation, is inefficient.

Unfortunately, this is not the world we live in. If combatants are unsure of how civilians respond to punishment, the balancing equilibrium breaks down and high levels of violence can emerge. Where territorial control is evenly divided between insurgents and the government, expectations of civilian bandwagoning create incentives for two-sided violence. Where one side has incomplete territorial control, violence by the weaker combatant is likely. Where one side has complete control, violence by the stronger combatant is likely.

These propositions differ from the emerging conventional wisdom (Kalyvas, 2006), which expects (a) violence in divided and fully-controlled areas to be off the equilibrium path, and (b) selective violence in areas of incomplete control to be perpetrated by the stronger combatant. Both of these expectations are built on the assumption that civilians always balance. The epidemic model shares Kalyvas' scope conditions and underlying logic – about the role of civilian cooperation in the production of selective violence, about the relationship between identification and territorial control, about the role of survival in civilian decision-making. However, the model accommodates both types of civilian behavior and shows that – once the balancing assumption is loosened – civilian victimization is not always counterproductive. Insurgents in a position of weakness have incentives to terrorize civilians. Violence between two equally-matched combatants can become a competitive killing spree.

Given the risks of escalation, why would civilians ever choose to bandwagon? Survival, as the epidemic model relates, is a shot in the dark. By the time civilians respond to the use or non-use of punishment, the blood will have already been spilt. Given the same level of violence, a civilian who balances bears the same costs as one who bandwagons. In making this choice, civilians do not directly control their own fate; they shape incentives for violence. How combatants respond to these incentives depends on whether civilian strategy can be clearly communicated, and whether the initial balance of power is symmetric, uneven or monopolistic. Where their strategy choice is not public knowledge, civilians have great difficulty realizing, much less exploiting their own leverage as kingmakers. In a virtual setting and empirically, civilians have trouble coalescing around a single, unified strategy and tend to realize the benefits of balancing when it is too late: after combatants have already begun to capitalize on bandwagoning through escalation. Neither rational cost-minimizing behavior nor an evolutionary process of trial-and-error is sufficient to shake this dynamic.

Absent a reliable enforcement mechanism, it is difficult to see how bandwagoning can be avoided on a group level. Our best hope may be to make its consequences less extreme. As simulations suggest, civilian casualties decline as one side consolidates its control and violence becomes more selective. If we are only interested in protecting civilians from their own worst instincts, the simplest policy solution may be to choose a side, help ensure its decisive victory, and let civilians find safety in the shadow of the Leviathan. Without improvements in intelligence gathering, strategic evaluation and outreach to the local community – the underlying challenges that make civilian strategy so difficult to discern and communicate – even such extreme solutions may prove insufficient. As a first step, we should acknowledge that bandwagoning is a frequent feature of civil conflict, and try to identify potential patterns of violence that we may otherwise overlook or underpredict.

Appendix A: Equilibrium stability analysis

The stability of the insurgent victory equilibrium in (8) can be shown through the linearization of the system in (2-4). Let \mathbf{J}_1 be the Jacobian matrix of the system evaluated at fixed point (8):

$$\mathbf{J}_1 = \begin{pmatrix} \frac{-k\mu_I}{u+\theta_G\rho_G} & -u - \theta_G\rho_G & \frac{-(u+\theta_G\rho_G)\mu_G}{\mu_I} \\ \frac{k\mu_I - (u+\theta_G\rho_G)(u+(1-\theta_G)\rho_G+(1-\theta_I)\rho_I)}{u+\theta_G\rho_G} & 0 & 0 \\ \frac{-k\mu_I}{u+\theta_G\rho_G} & 0 & \frac{(u+\theta_G\rho_G)\mu_G}{\mu_I} - u - \theta_I\rho_I \end{pmatrix} \quad (11)$$

with μ_I, μ_G as defined in (1). The equilibrium point (8) is stable if all the eigenvalues of \mathbf{J}_1 have negative real parts. In order for this to be true, we must have $\det(\mathbf{J}_1) > 0, \text{trace}(\mathbf{J}_1) < 0$. This condition holds as long as $R_I > R_G$ and any one of the following is true: (a) $\beta = 0, \rho_I = 0, \rho_G = 1, \theta_G = 1/2$, (b) $\beta = 0, \rho_I = 1, \rho_G = 1, \theta_G < 1/2$, (c) $\beta = 1, \rho_I = 1, \rho_G = 0, \theta_G = 1/2$, (d) $\beta = 1, \rho_I = 1, \rho_G = 1, \theta_G > 1/2$.

A similar approach can be used to prove the stability of the government victory equilibrium in (9). Let \mathbf{J}_2 be the Jacobian matrix of the system in (2-4) evaluated at fixed point (9):

$$\mathbf{J}_2 = \begin{pmatrix} \frac{-k\mu_G}{u+\theta_I\rho_I} & -\frac{(u+\theta_I\rho_I)\mu_I}{\mu_G} & -u - \theta_I\rho_I \\ 0 & \frac{(u+\theta_I\rho_I)\mu_I}{\mu_G} & -u - \theta_G\rho_G & 0 \\ \frac{-k\mu_G}{u+\theta_I\rho_I} & 0 & 0 & 0 \end{pmatrix} \quad (12)$$

The equilibrium point (9) is stable if $\det(\mathbf{J}_2) > 0, \text{trace}(\mathbf{J}_2) < 0$. This condition holds as long as $R_I < R_G$ and any one of the following is true: (a) $\beta = 0, \rho_I = 1, \rho_G = 0, \theta_G = 1/2$, (b) $\beta = 0, \rho_I = 1, \rho_G = 1, \theta_G > 1/2$, (c) $\beta = 1, \rho_I = 0, \rho_G = 1, \theta_G = 1/2$, (d) $\beta = 1, \rho_I = 1, \rho_G = 1, \theta_G < 1/2$.

Appendix B: Archival abbreviations

GARF: State Archives of the Russian Federation, Moscow.

RGVA: Russian State Military Archive, Moscow.

TsDAGOU: Central State Archive of Public Organizations of Ukraine, Kyiv.

F: file (*fond*); Op: catalog (*opis'*); D: case (*delo*); L: page (*list*).

References

- Abrahms, M. (2006). Why Terrorism Does Not Work. International Security 31(2).
- Arreguín-Toft, I. (2001). How the Weak Win Wars: A Theory of Asymmetric Conflict. International Security 26(1).
- Arreguín-Toft, I. (2003). The [F]utility of Barbarism: assessing the impact of the systematic harm of non-combatants in war. Paper presented at the annual convention of the American Political Science Association, Philadelphia, Pennsylvania.
- Balcells, L. (2010). Rivalry and Revenge. Violence against Civilians in Conventional Civil Wars. International Studies Quarterly 54(2), 291–313.
- Balcells, L. (2011). Continuation of politics by two means: direct and indirect violence in civil war. Journal of Conflict Resolution 55(3), 397–422.
- Braithwaite, A. (2010). Resisting infection: How state capacity conditions conflict contagion. Journal of Peace Research 47(3).
- Buhaug, H. and K. S. Gleditsch (2008). Contagion or confusion? Why conflicts cluster in space. International Studies Quarterly 52(2).
- Burds, J. (1997). AGENTURA: Soviet Informants’ Networks & the Ukrainian Rebel Underground in Galicia, 1944-1948. East European Politics and Societies 11(1), 89–130.
- Cáceres A., J. F. (2003). Basic Computational Exercises in Game Theory . Department of Economics, USACH.
- Carey, S. C. (2006). The Dynamic Relationship Between Protest and Repression. Political Research Quarterly 59(1).
- Christia, F. (2008). The Closest of Enemies: Alliance Formation in the Afghan and Bosnian Civil Wars . Ph.D. dissertation, Harvard University.
- Cleveland, R. B., W. S. Cleveland, J. E. McRae, and I. Terpenning (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. Journal of Official Statistics 6(1), 3–73.
- Crescenzi, M. (1999). Violence and Uncertainty in Transitions. Journal of Conflict Resolution 43.
- Dallin, A. (1981). German Rule in Russia. London: Macmillan.
- Downes, A. B. (2006). Hypotheses on the Effectiveness of Civilian Victimization in War. Paper prepared for the Annual Meeting of the International Studies Association, March 22-25.
- Downes, A. B. (2007). Draining the Sea by Filling the Graves: Investigating the Efficacy of Indiscriminate Violence as a Counterinsurgency Strategy .
- Dyukov, A., D. Valieva, O. Dranitsina, and M. Mints (Eds.) (2009). Povsednevnost’ terrora: Deyatel’nost’ natsionalisticheskikh formirovaniy v zapadnykh regionakh SSSR [Routine terror: The activities of nationalist formations in the western regions of the USSR], Volume 1. Moscow: Fond Istoricheskaya pamyat’.

- Elton, G. (1975). Taxation for War and Peace in Early-Tudor England. In J. Winter (Ed.), War and Economic Development: Essays in Memory of David Joslin, pp. 33–48. Cambridge University Press.
- Epstein, J. M. (1997). Nonlinear Dynamics, Mathematical Biology, and Social Science. The Santa Fe Institute / Addison-Wesley.
- Epstein, J. M. and R. L. Axtell (1996). Growing Artificial Societies: Social Science From the Bottom Up. Brookings Institution.
- Epstein, J. M., J. Parker, D. Cummings, and R. A. Hammond (2008). Coupled Contagion Dynamics of Fear and Disease: Mathematical and Computational Explorations. PLoS One 3(12).
- Esteban, J., M. Morelli, and D. Rohner (2010). Strategic Mass Killings. Economics Working Papers ECO2010/23, European University Institute.
- Francisco, R. (2004). After the massacre: Mobilization in the wake of harsh repression. Mobilization: An International Journal 9(2), 107–126.
- Gintis, H. (2009). Game Theory Evolving (2nd ed.). Princeton University Press.
- Goodwin, J. (2001). No Other Way Out: States and Revolutionary Movements, 1945-1991. Cambridge: Cambridge University Press.
- Gregory, P. R., P. J. Schröder, and K. Sonin (2011). Rational dictators and the killing of innocents: Data from Stalin’s archives. Journal of Comparative Economic 39, 34–42.
- Headquarters, Department of the Army (2006, 15 December). Counterinsurgency, Field Manual No. 3-24. Washington, DC: Headquarters, Department of the Army.
- Hehn, P. N. (1979). The German Struggle Against Yugoslav Guerrillas in World War II: German Counter-Insurgency in Yugoslavia, 1941-1943. Boulder, CO: Eastern European Quarterly.
- Hibbs, D. (1973). Mass Political Violence. John Wiley.
- Hultman, L. (2007). Battle losses and rebel violence: Raising the costs for fighting. Terrorism and Political Violence 19(2), 205–222.
- Kalyvas, S. (2008a). Ethnic Defection in Civil War. Comparative Political Studies 41(8), 1043–1068.
- Kalyvas, S. N. (1999). Wanton and Senseless? The Logic of Violence in Civil Wars. Rationality and Society 11(3).
- Kalyvas, S. N. (2004). The Logic of Terrorism in Civil War. Journal of Ethics 8(1).
- Kalyvas, S. N. (2005). Warfare in Civil Wars. In I. Duyvesteyn and J. Angstrom (Eds.), Rethinking the Nature of War.
- Kalyvas, S. N. (2006). The Logic of Violence in Civil War. Cambridge University Press.
- Kalyvas, S. N. (2008b). Promises and Pitfalls of an Emerging Research Program: The Microdynamics of Civil War. In S. N. Kalyvas, I. Shapiro, and T. Masoud (Eds.), Order, Conflict, Violence, pp. 1–14. Cambridge University Press.

- Kalyvas, S. N. and M. A. Kocher (2007). How 'Free' Is Free-Riding in Civil War? Violence, Insurgency, and the Collective Action Problem. World Politics 59(2).
- Kocher, M., T. Pepinsky, and S. Kalyvas (2011). Aerial Bombing and Counterinsurgency in the Vietnam War. American Journal of Political Science 55(2).
- Kolesin, I. D. (1998). Matematicheskaya model' mezhetnicheskikh otnosheniy [Mathematical model of ethnic relations]. Izvestiya AN. Ser. TiSU (3).
- Kolesin, I. D. (2000). Germanskiy natsionalizm kak sotsial'no-psikhologicheskii fenomen [German nationalism as a social-psychological phenomenon]. Psikhologicheskii zhurnal 21(3).
- Kolesin, I. D. (2010). Matematicheskie modeli sotsial'noy samoorganizatsii [Mathematical models of social self-organization]. Ph.D. Dissertation, St. Petersburg State University.
- Koval, M. V. (2003). Organizatsiya ukrainskikh natsionalistov (OUN): Uroki istorii [Organisation of Ukrainian Nationalists (OUN): Lessons of History]. Otechestvennaia istoria 1(31 Jan), 53–77.
- Kremer, M. (1996). Integrating Behavioral Choice into Epidemiological Models of AIDS. The Quarterly Journal of Economics 111(2).
- Lichbach, M. I. (1987). Deterrence or Escalation? : The Puzzle of Aggregate Studies of Repression and Dissent. Journal of Conflict Resolution 31.
- Lskavyan, V. (2007). A rational choice explanation for Stalin's 'Great Terror'. Economics & Politics 19(2).
- Lyall, J. (2009). Does Indiscriminate Violence Incite Insurgent Attacks? Evidence from Chechnya. Journal of Conflict Resolution 53(2).
- Lyall, J. (2010). Are Coethnics More Effective Counterinsurgents? Evidence from the Second Chechen War. American Political Science Review.
- Mason, T. D. and D. A. Krane (1989). The Political Economy of Death Squads: Toward a Theory of the Impact of State-Sanctioned Terror. International Studies Quarterly 33(2).
- May, R. M. and M. A. Nowak (1995). Coinfection and the Evolution of Parasite Virulence. Proceedings: Biological Sciences 261(1361), 209–215.
- Maynard Smith, J. (1982). Evolution and the Theory of Games. Cambridge, UK: Cambridge University Press.
- McKelvey, R. D. (1991). A Liapunov function for Nash equilibria. California Institute of Technology.
- McKelvey, R. D. and A. McLennan (1996). Computation of equilibria in finite games. In H. Amman and J. R. D. Kendrick (Eds.), Handbook of Computational Economics.
- Moore, W. H. (1998). Repression and Dissent: Substitution, Context, and Timing. American Journal of Political Science 42(3).
- Moore, W. H. (2000). The repression of dissent : A substitution model of government coercion. Journal of Conflict Resolution 44.
- Nowak, M. (2006). Evolutionary Dynamics. Cambridge, MA: Belknap.

- Nowak, M. A. and R. M. May (1994). Superinfection and the Evolution of Parasite Virulence. Proceedings: Biological Sciences 255(1342), 81–89.
- Nowak, M. A. and R. M. May (2000). Virus dynamics: mathematical principles of immunology and virology. Oxford University Press.
- Pape, R. A. (1996). Bombing to Win: Air Power and Coercion in War. Ithaca: Cornell University Press.
- Peters, R. (2007). Progress and peril: New counterinsurgency manual cheats on the history exam . Armed Forces Journal (144).
- Pierskala, J. H. (2010). Protest, deterrence, and escalation: The strategic calculus of government repression. Journal of Conflict Resolution 54(1).
- Presidium of Supreme Soviet of USSR, Information-Statistical Division (Ed.) (1946). SSSR: Administrativno-territorial'noye delenie soyuznykh respublik [USSR: Administrative division of Union Republics]. Moscow: Izd. 'Izvestiya Sovetov Deputatov Trudyashchikhsya SSSR'.
- Saxton, G. and M. Benson (2008). Means, motives and opportunities in ethno-nationalist mobilization. International Interactions 34(1), 53–83.
- Siverson, R. M. and H. Starr (1991). The Diffusion of War: A Study of Opportunity and Willingness. Ann Arbor: The University of Michigan Press.
- Statiev, A. (2005). Motivations and Goals of Soviet Deportations in the Western Borderlands. The Journal of Strategic Studies 28(6), 977–1003.
- Statiev, A. (2008). Was Smuglianka a Lunatic or a Siguranta's Agent-Provocateur? Peculiarities of the Soviet Partisan Struggle in the Western Borderlands. Journal of Strategic Studies 31(5), 743–770.
- Statiev, A. (2010). The Soviet Counterinsurgency in the Western Borderlands. Cambridge University Press.
- Stoll, D. (1993). Between Two Armies in the Ixil Towns of Guatemala. New York: Columbia University Press.
- Tilly, C. (1978). From Mobilization to Revolution. Addison-Wesley.
- Tilly, C. (1985). War making and state making as organized crime. In P. Evans and T. S. D. Rueschemeyer (Eds.), Bringing the State Back In, pp. 169–191. Cambridge.
- Tys-Krokhmaliuk, Y. (1972). UPA Warfare in Ukraine: Strategical, Tactical, and Organizational Problems of Ukrainian Resistance in World War II. New York: Vantage Press.
- Ungureanu, V. and A. Botnari (2005). Nash equilibria sets in mixed extension of 2x2x2 games. Computer Science Journal of Moldova 13(1 (37)), 13–28.
- Valentino, B., P. Huth, and D. Balch-Lindsay (2004). 'Draining the Sea': Mass Killing and Guerrilla Warfare. International Organization 58(2).
- Vladimirtsev, N. and A. Kokurin (2008). NKVD-MVD SSSR v bor'be s banditizmom i vooruzhennym natsionalisticheskim podpol'yem na Zapadnoy Ukraine, v Zapadnoy Belorussii i Pribaltike 1939-1956 [NKVD-MVD in the struggle against banditry and armed nationalist guerillas in Western Ukraine, Western Belarus and the Baltics]. Moscow: MVD.

- Wood, E. (2009). Armed groups and sexual violence: when is wartime rape rare? Politics and Society 37(1), 131–162.
- Zhukov, Y. M. (2012). Roads and the Diffusion of Insurgent Violence: The Logistics of Conflict in Russia's North Caucasus. Political Geography Forthcoming.